REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
4/15/96

3. REPORT TYPE AND DATES COVERED
12/1/92 to 2/29/96. FINAL

4. TITLE AND SUBTITLE
Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science

5. FUNDING NUMBERS
N00014-93-0152
Fund 22738A

6. AUTHOR(S)
S.B. Baden

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, San Diego
Computer Science & Engineering Dept.
9500 Gilman Drive
La Jolla, CA 92093-0114

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dept. of the Navy
Office of Naval Research
San Diego Regional Office
4520 Executive Drive, Suite 300
San Diego, CA 92121-3019

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Public Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The objectives of this investigation are to develop a new programming model, called lattice parallelism, that will significantly reduce the amount of effort required to parallelize non-uniform structured applications. We implemented the model as two C++ class libraries, called KeLP and LPARX. We used KeLP and LPARX to develop a real-space adaptive mesh solver for a significant grand challenge application---first principles simulations of real materials. Our adaptive solver reduces memory and time consumption by two orders of magnitude over a uniform mesh based method. The code is fully portable and runs on the Intel Paragon, Cray C90 and IBM SP2. Performance on the MPP systems is competitive with the Cray C90.

14. SUBJECT TERMS
Parallel Programming, Multiprocessors, Ab-initio Molecular Dynamics, Irregular Computations, Run Time Support

15. NUMBER OF PAGES
13

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

Final Report
April 11, 1996

Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science

Contract N00014-93-1-0152

Scott B. Baden
Department of Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093-0114

(619) 534-8861
baden@gili.ucsd.edu

Principal Investigator:	Scott B. Baden
Institution:	University of California, San Diego
Phone number:	(619) 534-8861
E-mail address:	baden@cs.ucsd.edu
Contract title:	Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science
Contract number:	N00014-93-1-0152

1 Numerical productivity measures

Refereed articles in preparation: 1
Refereed papers accepted or published: 14
Unrefereed reports and articles: 3
Invited Presentations: 27
PhD's Graduated: Scott R. Kohn (Ph.D. '95)
Graduate students supported >= 25 % of full time: 2

Principal Investigator:	Scott B. Baden
Institution:	University of California, San Diego
Phone number:	(619) 534-8861
E-mail address:	baden@cs.ucsd.edu
Contract title:	Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science
Contract number:	N00014-93-1-0152

2 Summary of technical results

2.1 Technical Objectives

The objectives of this investigation are to develop a new programming model, called lattice parallelism, that will significantly reduce the amount of effort required to parallelize non-uniform structured applications, and to demonstrate the effectiveness of lattice parallelism in a significant grand challenge application — first principles simulations of real materials.

We have chosen this particular application for two reasons. First, it encompasses a diversity of non-uniform structured numerical techniques which apply more generally to a wide spectrum of fields outside numerical science, such as computational fluid dynamics. Second, while first principles simulations have become highly accurate, it is still not possible to treat most problems of technological significance because of computational bottlenecks. Thus, our investigation presents a real opportunity to advance the state-of-the-art in parallel programming and to increase our knowledge about diverse physical phenomena.

We assert that the use of non-uniform structured methods in combination with parallel computers will provide a significant breakthrough in computational efficiency. First principles simulations are broadly applicable to many areas of materials science, such as the prediction of high temperature super-conductors, the understanding of catalytic reactions, the enhancement of diamond film growth. They have a large potential for improving our understanding of technologically important advanced materials.

An important aspect of our investigation is that it entails significant interdisciplinary research. We will build experimental software tools and employ them in experimental applications software. We are working closely with specialists in materials science, applied mathematics, and computational fluid dynamics, both within the academic community and at the national laboratories. This close collaboration has already produced results, both in terms of new a software infrastructure and new numerical approaches.

2.2 Technical Approach

2.2.1 Overview

We developed the prototype lattice parallelism system, called LPAR, and subsequently developed a production version called LPARX (Lattice PARallelism eXtended version.) LPARX has since been extended and redesigned, resulting in our latest system called KeLP (Kernel Lattice Parallelism.) These systems are domain-specific C++ class libraries targeted to irregular problems such as multilevel adaptive finite difference methods and particle methods. They reduce software development time by providing a high level machine-independent model that suppresses low-level details such as message passing. Such information hiding is useful in reducing software development costs because it promotes software portability, whereby software performance is robust with respect to changes in physical machine design. In addition, information hiding promotes software reuse: existing application software originally designed for a conventional single processor computer may be incorporated into parallelized applications with little or no change.

Portability is important both because parallel processing hardware is changing rapidly, and because parallel processors come in different varieties. LPARX and KeLP run on MPPs such as the Intel Paragon, IBM SP2, and Cray T3D, as well as workstation clusters running under PVM.

2.2.2 The Lattice Parallel Programming Model

LPARX and KeLP are alternative implementations of the Lattice Parallelism programming model. This model supports a coarse grain data parallelism, that is, the illusion of a single global address space and a single logical thread of control. A useful side effect of this property is that "parallel" applications can be conveniently developed on a single processor workstation, which is convenient for the applications programmer.

A unique aspect of lattice parallelism is its support for "structural abstraction." Structural abstraction enables the programmer to explicitly represent information about the structure and layout of distributed data separately from the data itself. Such "meta-data" is treated as a "first class" language object; it may be assigned storage and manipulated by the programmer much as ordinary numbers are treated in most programming languages. This has two major benefits. First, the meaning of data coordination is unaffected by the geometry of the data decomposition scheme, which permits the programmer to specify arbitrary data blocking schemes tuned to the application. Second, geometric set operations may be used to eliminate much of the bookkeeping traditionally required to manage communication. Applications written in LPARX and KeLP, such as adaptive mesh refinement, are typically much shorter than the equivalent fortran 77 encodings.

In theory, structural abstraction applies to arbitrary geometries, though in practice we restrict data layouts to finite unions of rectangles in an arbitrary d-dimensional space Z^d, where the rectangles are generally of different sizes. These decompositions are appropriate both for nested refinement structures arising in structured adaptive mesh methods and for particle methods that employ spatial decomposition. By comparison, languages such as Fortran 90 and HPF do not support run time layouts of data comprising non-uniformly sized blocks.

Structural abstraction also supports the development of dimension-independent application software. For example, 3d dimensional applications can be partially debugged—e.g. data structure management and load balancing—on a 2d version of the 3d problem. The time for a bug to manifest itself is reduced substantially, since 2d numerical calculations generally run more quickly than their 3d counterparts. These scaled-down computations can be carried out on a workstation rather than on production hardware, and the user has confidence that the code will work on the production hardware (ignoring differences in machine arithmetic) because the LPARX and KeLP abstractions are portable.

2.2.3 LPARX

LPARX provides dynamic memory management and data decomposition facilities that are unavailable in data parallel Fortran languages such as HPF. LPARX extends Fortran's simple notion of an array as a static, flat, rectilinear structure with the more flexible notion of dynamic structures comprising irregular hierarchical collections of arrays. Objects communicate within a shared name space, obviating the need to pass messages, and the user is able to manage locality within the memory hierarchy to avoid high overhead costs using the block copy operation.

2.2.4 KeLP

Optimizing communication is important on parallel computers due to the potentially large penalty associated with accessing off-processor data. Structural abstraction assists in reducing the overhead in communication, through "orchestrated" communication. It enables each processor to determine precise communication requirements using local information only. This is possible because the structural information is compact and may be efficiently replicated on all processors. Orchestration enables each processor to preallocate message receive buffers and to avoid unnecessary memory copying within the message passing subsystem and is related to the inspector/executor model developed by Saltz and coworkers. KeLP system supports orchestrated communication and was developed jointly with Ph.D. student S. Fink. KeLP's orchestrated communication has been observed to reduce the communication overheads in adaptive mesh methods, for example, by as much as a factor of 2 to 4 on the IBM SP2, as compared with LPARX, which does not support orchestrated communication.

KeLP's orchestrated communication relies on custom schedules, which extend the scheduling mechanisms provided by Multiblock PARTI, developed by Saltz and coworkers. However, unlike Multiblock PARTI, KeLP schedules are not opaque objects, over which the programmer has limited control. KeLP's custom schedules enable the programmer to define specialized schedules need to express irregular communication patterns arising in structured adaptive methods

2.2.5 Application Programmer Interfaces

We implemented an Application Programmer Interface (API) in LPARX and later ported the API to KeLP. The API enables the application programmer to focus on the mathematical aspects of the problem while remaining aloof of the low level details. The API provides services germane to adaptive mesh refinement algorithms—grid generation, grid hierarchy management and display—plus problem-specific parts that provide error estimation and smoothers.

The API has two requirements: it must be extensible and it must provide portable performance. Since our adaptive eigensolver is still under development, we may need to explore higher-order stencils. Such changes must be relatively easy to program. One way of meeting this requirement is to provide implementation layers like KeLP and LPARX that raise the level of detail seen by the API developer. This approach decouples the API from inevitable changes in hardware. Casual users will identify with the API. Sophisticated users may customize the API to the application at hand or write a new one.

We have observed that LPARX and KeLP applications achieve performance comparable to applications hand-coded with explicit message passing (i.e. with MPI), are simpler and easier to maintain.

2.2.6 User Base

LPARX and KeLP are employed in real applications within my research group, within other Departments at UCSD, and also outside UCSD. I use LPARX in graduate instruction. Students who have taken my graduate course have continued to use LPARX to carry out computations in their dissertation research, including Ph. D. student K. Zhang who is modeling micro-magnetics in the Physics Department (advisor: D. Fredkin) and Dr. W. Hart who is implementing Genetics Algorithms (Sandia National Laboratory; UCSD advisor: R. K. Belew.) Students S. Fink and C. Huston used LPARX to implement a dimension-independent code for connected component labeling for spin models in statistical mechanics. S. Figueira used LPARX to analyze performance tradeoffs of various parallelization strategies for localized N-body solvers. S. Kohn implemented a 3D smoothed particle hydrodynamics code for modeling the evolution of galaxy clusters in collaboration with with John F. Wallen (Institute for Computational Sciences and Informatics at George Mason University.)

KeLP and LPARX users outside UCSD include:

1. Chris Myers (Cornell Theory Center) has implemented a parallel 2D code to study localized slip modes in the dynamics of earthquake faults and has collaborated with J. Sethna (Cornell University) on a parallel code to study shape-memory effects in martensitic alloys. Myers's reported on experiences with LPARX at the 1995 meeting of the APS Physics Computing Conference ("Some ABCs of OOP for PDEs on MPPs") and also in a Cornell Theory Center "Smart Node" newsletter available via the World Wide Web at URL: http://www.tc.cornell.edu/SmartNodes/Newsletters/1994/V6n6/Myers.

2. G. Cook (Cornell University) has used LPARX to construct an API for adaptive multigrid methods in numerical relatively as part of the Black Hole Binary Grand Challenge Project.

3. G. Duncan (Bowling Green State University) has investigated the use of LPARX to parallelize an adaptive hyperbolic solver for simulations of relativistic jets.

4. T. Maxwell (University of Maryland Institute for Ecological Economics, Solomons) has used LPARX in spatio-temporal modeling of ecosystems.

5. The Center for Computational Sciences and Engineering (CCSE) at Lawrence Berkeley National Laboratory is employing orchestrated communication techniques pioneered in KeLP to implement their own adaptive mesh refinement class libraries for gas dynamics on parallel processors.

2.2.7 Computational Accomplishments

The collaboration with Weare et al. has enabled us to tackle a difficult eigenvalue problem in materials science that currently cannot be addressed by current methods. We have developed an adaptive multigrid eigenvalue solver (AES) that employs structured refinements.

AES has been applied to simple molecules. Our solver can compute the potential of the H2+ molecule to greater accuracy than that attainable with traditional FFT-based methods. Performance of our AMG solver on the Intel Paragon is competitive with the Cray C-90 (single CPU): a computation involving H2+ (described below) runs in 57 seconds on 8 processors of the Paragon and in 28 seconds on the C90. Adaptivity is essential in reducing computation time and memory overheads by TWO ORDERS of magnitude.

We are currently investigating more elaborate molecules that exhibit several length scales. For example the HHe++ molecule which has a relatively short length scale for the He atom with charge 2++ and a longer scale for the H atom with a single charge. Problems such as this are typical of the application of these approaches to real heterogeneous materials and pose a significant problem for uniform grid methods which have to resolve the smallest length as well as assure complete coverage for the longer length.

2.2.8 Importance of the Accomplishments

The potential technological impact of our investigation is concentrated in two key areas:

1. software infrastructure for parallel adaptive methods;

2. first principles simulation carbon fibers and other complex materials.

The Application Programmer Interfaces developed by our group may be applied to a number of scientific and engineering disciplines, and thus our research will have an impact in other fields such as computational fluid dynamics and semiconductor modeling. The orchestration mechanisms explored by KeLP are unique in that they employ geometric primitives to express communication. This model has applications to diverse data motion problems including: data intensive applications, involving large volumes of I/O between memory, disk, and tape; multidisciplinary applications, which couple two or more separate applications; out of core applications, where data is too large to fit into memory and must be moved in stages to and from disk; and hierarchical parallelism, in which each processor is in fact a parallel processor.

Our investigation of adaptive non-linear eigensolvers may lead to an improved understanding of technologically important materials such as carbon fibers. More generally, the parallel adaptive numerical techniques we propose may broaden the range of materials that can be studied at a quantitative predictive level that currently cannot be treated with existing (uniform) methods.

Principal Investigator:	Scott B. Baden
Institution:	University of California, San Diego
Phone number:	(619) 534-8861
E-mail address:	baden@cs.ucsd.edu
Contract title:	Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science
Contract number:	N00014-93-1-0152

3 Lists of publications, presentations and reports

Publications

Journals

1. J. R. Pilkington and S. B. Baden, "Dynamic Partitioning of Non-Uniform Structured Workloads with Spacefilling Curves," *IEEE Transactions on Parallel and Distributed Computing,* 7(2), Februrary 1996. *In Press.*

2. S. R. Kohn and S. B. Baden, "Irregular Coarse-Grain Data Parallelism Under LPARX," *Journal of Scientific Programming,* 5(3), Fall 1996. *In Press.*

3. S. B. Baden and S. R. Kohn, "Portable Parallel Programming of Numerical Problems Under the LPAR System," *Journal of Parallel and Distributed Computing,* **27**, 1995, pp. 38–55.

CONFERENCES

1. W. E. Hart, S. B. Baden, R. K. Belew, and S. R. Kohn, "Analysis of the Numerical Effects of Parallelism on a Parallel Genetic Algorithm." 10th International Parallel Processing Symposium, Honolulu, HI, September 1995. *In Press.*

2. S. R. Kohn and S. B. Baden, "A Parallel Software Infrastructure for Structured Adaptive Mesh Methods," Supercomputing '95, December 4-7 1995, San Diego, California.

3. "Parallel Structured Adaptive Real-Space Methods for the Solution of the LDA Equations for Materials Applications," S. R. Kohn, R. Kawai, S. B. Baden, M. E. G. Ong, and J. H. Weare, Materials Research Society, November 1995, Boston, Massachusetts. (Abstract)

4. S. M. Figueira and S. B. Baden, "Performance Analysis of Parallel Strategies for Localized N-body Solvers," Seventh SIAM Conference on Parallel Processing for Scientific Computing, February 15-17, 1995, San Francisco, California, pp. 349–354.

5. S. J. Fink and S. B. Baden, "Run-time Data Distribution for Block-Structured Applications on Distributed Memory Computers," Seventh SIAM Conference on Parallel Processing for Scientific Computing, February 15-17, 1995, San Francisco, California, pp. 762–767.

6. S. R. Kohn and S. B. Baden, "The Parallelization of an Adaptive Multigrid Eigenvalue Solver with LPARX," Seventh SIAM Conference on Parallel Processing for Scientific Computing, February 15-17, 1995, San Francisco, California, pp. 552–557.

7. E. J. Bylaska, S. R. Kohn, S. B. Baden A. Edelman, R. Kawai, M. E. G. Ong, and J. H. Weare, "Scalable Parallel Numerical Methods and Software Tools for Material Design," Seventh

SIAM Conference on Parallel Processing for Scientific Computing, February 15-17, 1995, San Francisco, California, pp. 219–224.

8. S. B. Baden and S. R. Kohn and S. J. Fink, "Programming in LPARX," in *Proceedings of 1994 Intel Supercomputer User's Group Conference*, San Diego, California, June 25-29, 1994, IEEE.

9. S. J. Fink, S. B. Baden, and K. Jansen, "Cluster Identification on a Distributed Memory Multiprocessor," in *Proceedings of the 1994 Scalable High Performance Computing Conference*, Knoxville, Tennessee, May 23-25, 1994, IEEE, pp. 239–246.

10. S. R. Kohn and S. B. Baden, "A Robust Parallel Programming Model for Dynamic Non-Uniform Scientific Computations," in *Proceedings of the 1994 Scalable High Performance Computing Conference*, Knoxville, Tennessee, May 23-25, 1994, IEEE, pp. 509–517.

11. S. R. Kohn and S. B. Baden, "An Implementation of the LPAR Parallel Programming Model for Scientific Computations," in *Proceedings of the Sixth SIAM Conference on Parallel Processing for Scientific Computing*, Norfolk, Virginia, March 22-24, 1993, pp. 759–766.

12. S. J. Fink, C. Huston and S. B. Baden, "Parallel Cluster Identification for Multidimensional Lattices," submitted for publication to *IEEE Transactions on Parallel and Distributed Computing,* 1995.

Presentations

- *Software Infrastructure for Irregular Scientific Computations on Parallel Processors.* Computer Science Division, University of California, Berkeley, November 1995. Applied Mathematics Department, California Institute of Technology, Pasadena, California, October 1995. SoCal Parallel Computing Meeting, Pasadena, California, September 1995. GDE Systems, Rancho Bernardo, California, September 1995.

  Computational Alchemy Using Ab-initio Molecular Dynamics: Computational Challenges, Technological Rewards. PetaFLOPS Summer Study, Bodega Bay, California, August 1995.

- *The Computational Alchemist's Laboratory: From Atoms to Abstractions.* Lawrence Livermore National Laboratory, August 1995. Cornell Theory Center, Cornell University, August 1995. High Performance Computing Group, Digital Equipment, Corporation, Maynard, Massachusetts, August 1995. Department of Computer Sciences, The University of Texas at Austin, June 1995.

- *The LPARX System.* Seventh SIAM Conference on Parallel Processing for Scientific Computing, San Francisco, California, February 1995.

- *Digital Alchemy: From Atoms to Abstractions.* Center for Innovative Computer Applications, Indiana University, Bloomington, November 1994. Oak Ridge National Laboratory, November 1994. Argonne National Laboratory, Math and Computer Science Division, October 1994. Computer Sciences Department, University of Wisconsin, Madison, September 1994.

- *LPARX: An Approach to Implementing Multilevel Adaptive Methods on Parallel Computers.* Workshop in Materials Science: *Ab Initio* Local Density Approximations for Electronic Structure Calculations, San Diego, California, July 1994, U.C. San Diego, San Diego Supercomputer Center, SIAM.

- *Portable Scientific Programming under LPARX.* SIAM Annual Meeting, San Diego, California, July 1994.

- *Coarse Grained Data Parallel Programming with LPARX.* Center for Research in Parallel Computation, Rice University, Houston, Texas, June 1994. Texas Center for Applied Molecular Computation, University of Houston, Houston, Texas, June 1994.

- *Portable Scientific Programming.* Los Alamos National Laboratory, Los Alamos, New Mexico, January 1994. Department of Computer Science and Engineering, University of Washington, Seattle, November, 1993.

- *Programming Scientific Calculations with the LPAR Programming System.* Lawrence Livermore National Laboratory, August 1993. MetaCenter Computational Science Institute in Parallel Computing, San Diego Supercomputer Center, August 1993.

- *Programming Scientific Calculations with LPAR.* Computer Sciences Department, University of Wisconsin, Madison, June 1993.

- *The LPAR Programming System for Parallel Computation of Scientific Problems.* Computer Science Division, University of California, Berkeley, April 1993. Computer Science Department, University of Maryland, College Park, March 1993. National Science Foundation, Washington, D.C., March 1993.

- *Considerations in Using Massively Parallel Computers: A Minimalist Approach to Writing Robust Software,* UCSD Connect "Meet the Researcher Series," San Diego Supercomputer Center, January 1993.

- *The Role of Heuristics in Parallel Computation of Scientific Problems.* Three Decades of Numerical Linear Algebra at Berkeley: A Conference in Honor of the Sixtieth Birthdays of Beresford Parlett and William Kahan, Berkeley, California, October 1992.

Principal Investigator:	Scott B. Baden
Institution:	University of California, San Diego
Phone number:	(619) 534-8861
E-mail address:	baden@cs.ucsd.edu
Contract title:	Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science
Contract number:	N00014-93-1-0152

4 Description of research transitions and DoD interactions

We are continuing to interact with Dr. Charles Rendleman of the CCSE Group at Lawrence Berkeley National Laboratory. The LBNL group is adopting techniques developed in LPARX and KeLP to reduce the proliferation of architecture-dependent code streams, and to optimize communication overheads on MPPs.

We also interact with scientists at Los Alamos National Laboratory, in particular Dan Quinlan. Dr. Quinlan is developing an array class library, called P++, for adaptive mesh methods. Together, KeLP/LPARX and P++ are helpful in simplifying the implementation of complex, adaptive mesh refinement algorithms.

Principal Investigator:	Scott B. Baden
Institution:	University of California, San Diego
Phone number:	(619) 534-8861
E-mail address:	baden@cs.ucsd.edu
Contract title:	Parallel Programming Methodologies for Non-Uniform Structured Problems in Materials Science
Contract number:	N00014-93-1-0152

5 Description of software and hardware prototypes

Both KeLP and LPARX are publicly available via the World Wide Web. The LPARX distribution includes the adaptive mesh API along with with the real space adaptive multigrid solver. Software and documentation may be found at the following URLs:

PI:	http://www-cse.ucsd.edu/users/baden
LPARX Software Distribution:	http://www-cse.ucsd.edu/users/baden/lparx.html
KeLP Software Distribution:	http://www-cse.ucsd.edu/groups/hpcl/scg/kelp.html
Materials Science:	http://www-cse.ucsd.edu/users/baden/first.html